基于移不变字典学习和稀疏编码的滚动轴承故障识别算法 *

曲建岭 [1]，余　路 [1†]，高　峰 [1]，田沿平 [1]，李　俨 [2]

(1. 海军航空大学青岛校区，山东 青岛 266041; 2. 西北工业大学 自动化学院，西安 710072)

摘　要：针对现有旋转机械故障识别算法过度依赖专家先验知识的问题，提出了一种基于移不变字典学习和稀疏编码（SIDL-SC）的自适应故障识别算法。首先，将不同故障状态下的振动信号进行分段和平滑预处理以降低数据处理复杂度，接着将加入自适应惩罚因子的移不变字典学习算法用于提取不同故障状态下的移不变基函数；然后，利用高效的特征符号搜索算法求解待识别信号在不同基函数下的稀疏系数以实现对待识别信号的重构；最后，以重构残差作为对该信号故障状态识别的判断依据。滚动轴承振动数据库和实测航空发动机振动信号的实验结果表明，该算法相较于现有算法具有更高的故障识别准确率，在实际中具有较强的可行性。

关键词：移不变字典学习；稀疏编码；特征符号搜索；振动信号；故障识别

中图分类号：TP391.4　　doi: 10.3969/j.issn.1001-3695.2017.07.0687

Fault recognition algorithm for rolling bearings based on shift invariant dictionary learning and sparse coding

Qu Jianling[1], Yu Lu[1], Gao Feng[1], Tian Yanping[1], Li Yan[2]

(1. *Qingdao Branch of Naval Aviation University*, *Qingdao Shandong* 266041, *China*; 2. *School of Automation*, *Northwestern Polytechnical University*, *Xi'an* 710072, *China*)

Abstract: According to current algorithms for rotating machines largely depending on expert prior knowledge, the paper proposed an adaptive fault recognition algorithm based on shift invariant dictionary learning and sparse coding. Firstly, it segmented and smoothed vibration signals to decrease the complexity. Then, it used shift invariant dictionary learning with adaptive penalty factor to learn shift invariant bases in different fault states. After that, it used an efficient sparse coefficient solver called Feature Sign Search for reconstructing signal to be recognized. Lastly, residual was an evidence to determining fault state the signal belonging to. In the experiments of rolling bearing datasets and vibration signals of real aero-engine demonstrate its higher accuracy than up-to-date algorithms and feasibility for practical applications.

Key Words: shift invariant dictionary learning; sparse coding; feature sign search; vibration signal; fault diagnosis

0　引言

对机械设备进行状态监控和故障诊断对于提高生产效率、预防事故发生具有重要的实际意义。在机械故障诊断中，对监控设备有效提取故障特征是实现故障模式识别的重要一环。然而，机械设备多样的工况和复杂的机械结构给特征提取带来了难度，基于时域和频域的特征提取算法依赖于人工特征选择，往往需要相关领域内的专家知识才能较好的完成特征提取工作[1,2]。

一直以来，研究人员针对机械振动故障特征提取做了大量的工作。其中，Harmouche 等人[3]利用频谱分析法研究不同轴承故障下的特征频率和包络谱，通过构建线性分类器实现不同故障的识别。但是基于频谱的分析方法依然依赖于人工经验知识，缺乏对数据自身的有效挖掘。王录雁等人[4]提出了基于EMD（experience mode decomposition）和灰色关联理论的轴承故障诊断方法，利用 EMD 将振动信号进行内禀模态分量（IMF）分解，建立各 IMF 能量分布与轴承故障状态的灰色关联模型实现故障诊断。但是 EMD 自身存在模态混叠现象，对于信号不同状态下的能量分布可能会产生干扰。蔡剑华[5]等提出了基于高阶统计量的滚动轴承故障诊断方法，该方法的实质是将信号的高阶谱恢复为功率谱，利用功率谱提取信号的故障特征信息实现故障诊断。但是高阶统计量存在受奇异值影响较大的缺陷，

基金项目：国家自然科学基金资助项目（51505491）；航空科学基金资助项目（20165853040）

作者简介：曲建岭（1968-），男，山东莱阳人，教授，博导，博士，主要研究方向为模式识别、飞参数据处理；余路（通讯作者）（1989-），男，博士研究生，主要研究方向为模式识别、信号处理（yulu_china@163.com）；高峰（1967-），女，山东莱阳人，副教授，硕士，主要研究方向为飞参数据处理与应用；田沿平（1989-），女，吉林农安人，讲师，硕士，主要研究方向为模式识别、信号与信息处理；李俨（1973-），男，陕西西安人，教授，博导，博士，主要研究方向为控制理论与控制工程.

导致算法的结果并不稳定。

近年来基于字典学习的特征提取算法成为机器学习领域的研究热点，其核心思想是通过字典学习在原始数据中构建稀疏变换域，使待识别信号在该变换域下的投影具有稀疏性，Wright 等人[6]提出了基于稀疏表示分类（sparse representation classification，SRC）的人脸识别算法，首次在稀疏表示理论与模式识别之间建立了纽带，使基于稀疏表示的特征能够作为模式识别的特征向量实现对样本模式的准确识别，这给故障模式识别带来了新的思路。

字典的构建方式通常分为两种：基于数学分析的字典构建方法和自适应字典构建方法。基于数学分析的字典构建方法针对原始数据的先验知识实现稀疏化，常以正交变换基的形式实现过完备字典的构建。振动数据常以离散余弦变换[7]（DFT）和傅里叶变换[8]（FFT）作为正交变换基。使用上述这些正交变换基虽然具有相对完善的算法框架，但缺乏灵活性，不能保证对应的数据类型在该变换基下是最稀疏的表示，某些信号特别是时频变化范围较广的信号在上述变换基下仍不稀疏或不够稀疏，从而影响对数据的重构精度[9]。基于自适应的字典构建方法则以原始数据为训练样本，通过 K-奇异值分解[10]（K-SVD）等算法实现过完备字典的构造，往往采用这种方法构造的字典对于原始数据的表达更具有稀疏性。从模式识别角度来看，越稀疏的特征往往对应着更佳的分类效果。

对于旋转机械设备而言，同一特征的出现具有循环往复性，因此本文考虑将移不变字典学习算法引入旋转机械故障识别中，以期能从原始数据中捕捉到具有时移不变性质的本质特征构成基函数，将这些特征平移来构建过完备字典，实现对机械设备故障状态的有效识别。

1 SIDL-SC 故障识别算法理论

1.1 移不变特征自学习算法

2006 年，Smith[11]首先在《Nature》杂志上提出了移不变稀疏编码学习算法(shift invariant sparse coding，SISC)并成功应用于声信号特征提取。在形式上，SISC 理论认为同一特征模式在时间序列上的移动是一个独立事件，学习到的基函数具有相对时间的移不变性。在构成学习字典的原子上，SISC 学习到的基函数是信号中反复出现的本质特征，每个字典原子体现了信号的某一方面的特征。因此，SISC 良好的移不变特性使其已在图像处理[12]、信号处理[13]等领域得到了应用。

SISC 使用一个包含 M 个基函数的字典 $[d_1,d_2,...,d_M]\in\mathbb{R}^{Q\times M}$，对于任意输入信号 $y_i\in\mathbb{R}^N$ 均可由这些基函数和相应系数的卷积和来表示，基本数学模型为

$$y_i=\sum_{m=1}^{M}d_m * x_{m,k}+\varepsilon \tag{1}$$

其中：“$*$”表示卷积算子，用于表示基函数在所有可能位置上的移动。系数 $x_{m,k}\in\mathbb{R}^P$ 具有稀疏性，ε 通常表示高斯白噪声。各参数满足以下关系：$Q<N$，$P<N$ 且 $P+Q-1=N$。定义目标函数

$$\langle d,x\rangle=\min_{d,x}\sum_{i=1}^{K}\left\|y_i-\sum_{m=1}^{M}d_m * x_{m,k}\right\|_2^2+\beta\sum_{m,k}\left\|x_{m,k}\right\|_1 \tag{2}$$

其中：稀疏惩罚项 β 用于控制稀疏化程度。通常 β 设为固定值，但由于训练样本长度的不同，固定值往往并不能达到最优解，因此本文考虑对其进行改进以适应在不同训练样本长度下 β 能具有较好的控制稀疏化效果。令 $\beta=\theta\cdot L$，其中 L 为训练样本长度，θ 定义为稀疏化尺度（通常 $0<\theta<0.1$），由此实现对稀疏惩罚项的改进。

字典学习阶段若同时优化基函数 d_m 和系数 $x_{m,k}$ 则该优化问题为非凸的，无法得到稳定解。因此，借鉴稀疏编码中交替优化的思想，依次更新 d_m 和 $x_{m,k}$，反复迭代直到目标函数收敛。优化过程分为系数求解和字典学习两部分[14]。

在系数求解部分，固定字典 D 求解稀疏系数 x，则式(1)模型被分解为 $i=1,...,N$ 共 N 个独立的优化问题，针对每个输入信号 y_i 求解得到的系数都独立于其他的输入，因此考虑单个输入 y_i 的系数求解优化问题。

$$\min_{x}\left\|y_i-\sum_{j=1}^{M}d_j * x_j\right\|_2^2+\beta\sum_{j=1}^{M}\left\|x_j\right\|_1 \tag{3}$$

本文采用高效的 Batch-OMP 算法[15]，在字典 D 中寻找最佳的匹配原子 x，其表达式为

$$\{x^k\}=\arg\min_{\{x^k\}}\sum_{k}\left\|y_i-d_jx^k\right\|^2,s.t.\left\|\alpha^k\right\|_0\le l,\quad\forall k \tag{4}$$

在字典学习部分，固定系数 x 求解字典矩阵 D，则原优化问题转换为 L_2 约束优化问题：

$$\min_{d}\sum_{i=1}^{N}\left\|y_i-\sum_{j=1}^{M}d_j * x_{i,j}\right\|_2^2,s.t.\left\|d_j\right\|_2^2\le c,\ 1\le j\le M \tag{5}$$

直接求解上式计算量巨大，本文在字典学习阶段借鉴时频域转换的思想，将时域计算量较大的卷积运算转换到频域，由频域的乘积运算代替时域的卷积运算，使运算复杂度由 $O(N^2)$ 降低为 $O(N\log N)$，提高了运算效率。由 Parseval 定理可知，式(5)可转换为如下优化问题[14]：

$$\min_{\hat{d}}\sum_{i=1}^{N}\left\|\hat{y}_i-\sum_{j=1}^{M}\hat{d}_j\cdot\hat{x}_{i,j}\right\|_2^2,s.t.\left\|\hat{d}_j\right\|_2^2\le c \tag{6}$$

其中：c 为常数，$\hat{d}_j$ 为基函数 d_j 的离散傅里叶变换，相应地，$\hat{x}_{i,j}$ 和 $\hat{y}_i$ 分别为 $x_{i,j}$ 和 y_i 的离散傅里叶变换。使用拉格朗日乘子法求解该问题，则该问题被分解为二次项求和的形式，对于每一个频域上的频率 f，构造拉格朗日函数：

$$L(\hat{d},\lambda)=\sum_{f}\left(\left\|\hat{y}^{(f)}-\hat{x}^{(f)}\hat{d}^{(f)}\right\|_2^2+(\hat{d}^{(f)})^{*}\Lambda\hat{d}^{(f)}\right)-cl^T\lambda \tag{7}$$

其中：$\lambda\in\mathbb{R}^M$ 为对偶变量，单位向量 $l\in\mathbb{R}^M$，并且

$$\hat{d}_f=\begin{pmatrix}\hat{d}_1^{(f)}\\ \vdots\\ \hat{d}_M^{(f)}\end{pmatrix},\hat{y}_f=\begin{pmatrix}\hat{y}_1^{(f)}\\ \vdots\\ \hat{y}_N^{(f)}\end{pmatrix},\Lambda=diag(\lambda),\hat{x}_f=\begin{pmatrix}\hat{x}_{(1,1)}^{(f)} & \hat{x}_{(1,2)}^{(f)} & \cdots\\ \hat{x}_{(2,1)}^{(f)} & \hat{x}_{(2,2)}^{(f)} & \cdots\\ \vdots & \vdots & \ddots\end{pmatrix}$$

$L(\hat{d},\lambda)$ 是复数函数，通过复数公式 $|z|^2=|\mathrm{Re}(z)|^2+|\mathrm{Im}(z)|^2$ 将复数函数转换为实函数求解，将 $\mathrm{Re}(\hat{d})$ 和 $\mathrm{Im}(\hat{d})$ 计算目标函数 $\hat{d}_{\min}^{(f)}=\arg\min_{\hat{d}} L(\hat{d},\lambda)$，优化参数 λ 可利用牛顿法求解，代入式(8)即实现对基的迭代更新。

$$\hat{d}_{\min}^{(f)}=((\hat{x}^{(f)})^{*}\hat{x}^{(f)}+\Lambda)^{-1}(\hat{x}^{(f)})^{*}\hat{y}^{(f)} \tag{8}$$

1.2 基于特征符号搜索的系数求解算法

考虑到 Batch-OMP 算法需要预先设定稀疏度等参数，这在一定程度上增加了算法的计算量，如果参数选择不当很可能收敛速度较慢甚至难以收敛。考虑到待识别样本的稀疏度一般未知，因此本文在测试样本识别阶段引入特征符号搜索算法(Feature Sign Search，FS)求解待识别信号在不同基函数上的重构系数。FS 算法由 Lee 提出[16]，其核心思想是通过对系数符号的猜测将系数分量分为可行集（Φ）和 0 集（Θ），在 0 集中不断搜索梯度变化最大的分量并将其对应的下标 i 放入可行集中直到算法收敛，算法具体步骤如下：

a)初始化 $x=\vec{0}$，$\theta=\vec{0}$，其中 $\theta_i=\{-1,0,1\}$ 表示系数分量 x_i 的符号。

b)在系数为零的分量中搜索梯度变化最大的分量及其下标放入可行集中。

$$i=\arg\max_i\left|\frac{\partial\|y-Ax\|^2}{\partial x_i}\right|,\forall\theta_i=0,\Phi=\Phi\cup\{i\} \tag{9}$$

若 $\left|\frac{\partial\|y-Ax\|^2}{\partial x_i}\right|>\gamma$，则 $\theta_i=-1$；若 $\left|\frac{\partial\|y-Ax\|^2}{\partial x_i}\right|<-\gamma$，则 $\theta_i=1$。

c)选取矩阵 A 的子矩阵 $\hat{A}$，$\hat{A}$ 中只含有与可行集中 i 相对应的列向量，同样，选取与 i 相应的子系数向量 $\hat{x}$ 和 $\hat{\theta}$，计算无约束优化方程 $\min_{\hat{x}} f(x)\equiv\|y-\hat{A}\hat{x}\|^2+\gamma\hat{\theta}^T\hat{x}$，求得最优解：

$$\hat{x}_{new}:=(\hat{A}^T\hat{A})^{-1}(\hat{A}^T y-\gamma\hat{\theta}^T/2) \tag{10}$$

在 $\hat{x}$ 和 $\hat{x}_{new}$ 之间进行线搜索，找到所有符号改变的系数向量，并将它们的目标函数与在 $\hat{x}_{new}$ 的目标函数进行比较，将 $\hat{x}_{new}$ 修正为使得目标函数值最小的系数向量。将 $\hat{x}$ 中系数分量为零的 i 从可行集移除，并更新符号集合 θ。

d)(a)若 $\forall x_j\neq 0$，都有 $\frac{\partial\|y-Ax\|^2}{\partial x_i}+\gamma sign(x_j)=0$ 成立，则转条件(b)，否则转 c)；

(b)若 $\forall x_j=0$，都有 $\frac{\partial\|y-Ax\|^2}{\partial x_i}\leq\gamma$，则得到最优解 x，否则返回 b)。

1.3 基于重构残差的故障识别算法

理论上来说，与待识别信号的真实状态越接近，则由基函数及其求解的稀疏系数重构的信号与待识别信号的误差越小，因此考虑将重构残差作为故障识别的判断准则。利用 FS 算法求解待识别信号 y_0 在各个状态移不变基函数下的稀疏系数 $x_1,\dots,x_k$。由式(11)求解基函数及其对应的稀疏系数的重构 $\hat{y}_1,\dots,\hat{y}_k$。

$$\hat{y}=\sum_{j=1}^{M}d_j * x_j \tag{11}$$

因此，SIDL-SC 算法的实现分为数据预处理、移不变字典学习、重构系数求解和计算残差四部分，具体实现步骤如图 1 所示。

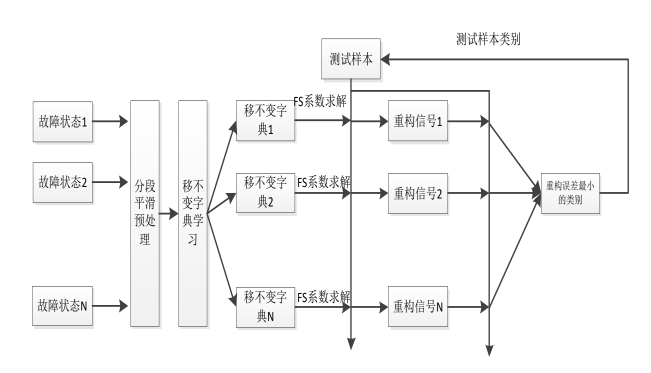


图 1　本文故障识别算法流程图

2　实验

2.1　实验设置

本文选取美国凯斯西储大学的开放轴承振动数据库进行实验[17]，使用电火花加工技术分别在轴承外圈、内圈和滚动体上设置单点故障，故障直径分别为 0.18 mm，0.36 mm，0.54 mm。本文将不同故障数据分为两组，如表 1 所示。其中实验 1 使用数据集 A 中的样本进行训练和测试，实验 2 中采用实验 1 的字典学习到的基函数对 1~2 hp 负载下的样本进行测试，以检测本文算法的泛化能力。

表 1　实验数据集描述

数据集	负载/hp	故障类型	故障尺寸/mm	类别标记
A/B	0/1~3	正常	0	1
		内圈故障	0.18	2
		内圈故障	0.36	3
		内圈故障	0.54	4
		外圈故障	0.18	5
		外圈故障	0.36	6
		外圈故障	0.54	7
		滚动体故障	0.18	8
		滚动体故障	0.36	9
		滚动体故障	0.54	10

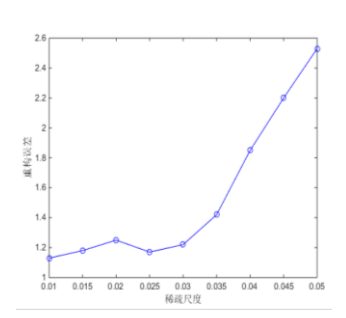
2.2　SIDL-SC 模型的建立

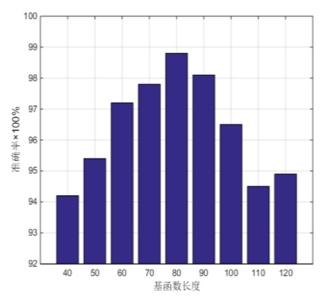
参数的选择影响算法的计算效率和故障识别的准确度，因此有必要根据训练样本的实际情况筛选出各个参数的最优值作为实现较高状态识别率的保证，本节考察对于建立 SIDL-SC 模

型起重要作用的几个参数。

2.2.1 稀疏化尺度

在移不变字典学习中，本文采用自适应稀疏尺度参数 θ 控制稀疏系数在目标函数中的比重，防止系数过大而基函数过小。实际表现为较大的 θ 值具有较快的收敛速度但是较高的重构误差。图 2(a)以实验集 A 中第 2 类内圈 18mm 故障样本为例，比较了不同 θ 值下移不变字典学习重构的原始信号的残差。从图中可以看出当稀疏尺度 θ 较低时，具有较低的重构误差，随着稀疏尺度的增大，重构误差呈现显著增大的趋势。但是从计算效率来说， θ 较小带来了每次迭代需要的计算量较大，从而导致收敛速度较慢。因此从重构误差和计算时间效率两方面来均衡，选择稀疏尺度为 0.025 是合适的。





(a)稀疏尺度对重构误差的影响 (b)基函数长度与识别准确率的关系

图 2 稀疏尺度和基函数长度对重构误差的影响

2.2.2 基函数长度的选择

在移不变字典学习中，基函数的长度的选择影响对样本的识别率。对于滚动轴承来说，学习到的基函数应该包含轴承中至少一个完整的冲击，从实际测量来看，滚动轴承的两个相邻波峰的间隔平均在 76 左右。同样以上一节中的实验样本为例，不同基函数长度下的样本识别率如图 2(b)所示，从图中可以看出，基函数在 70~90 能够达到较高的样本识别率。而基函数长度较短时，由于每个移不变基函数不能学习到完整的信号冲击，导致正确率不能达到最佳。而基函数长度较长可能会带来信息冗余，不利于后期使用 FS 进行重构。因此，本文选择基函数长度为 80。

2.2.3 训练样本长度的选择

在多次实验中发现不同的训练样本长度对于故障识别的作用很大。以 0 hp 负载下第 1、2、5、8 类样本为例进行训练和测试。其中每类故障状态随机选择 100 个样本用于测试，实验结果如图 3 所示。从图中可以明显看出，当训练样本长度为 128 和 256 时，在第 8 类故障的重构中，第 2、5、8 类学习到的基函数都能够以较低的残差对其重构，具体表现在这三类基函数的重构误差出现了混叠，这给样本的正确识别带来了困难。而输入样本为 512 和 1024 则较好的避免了混叠的现象，通过对比原始信号发现，训练样本长度为 128 和 256 分别包含 1 和 3 个完整冲击，而长度为 512 和 1024 包含有更多的完整冲击，从基函数自学习的角度来看，较多的冲击能使基函数学习到的特征更加显著，更有利于后期的重构分类。但训练样本长度的增加会带来每批次学习时间的增长，因此选择训练样本长度为 512。

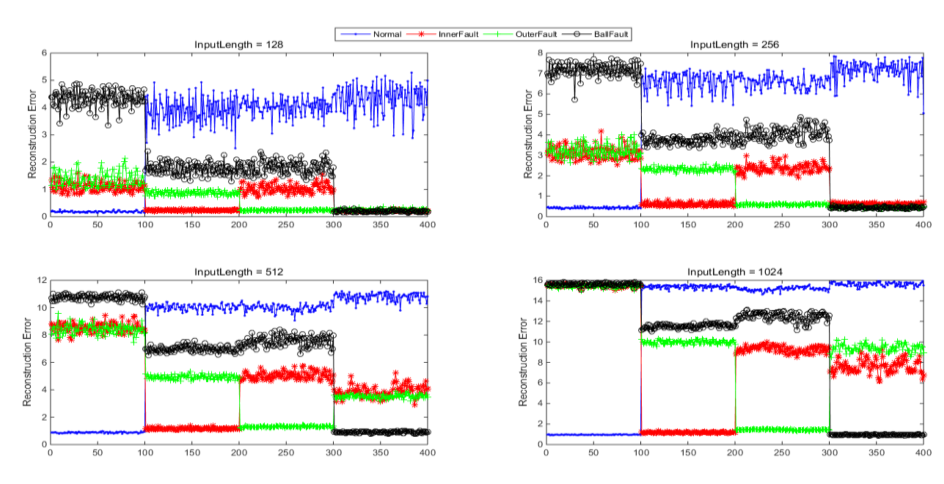


图 3 训练样本长度与重构误差的关系

2.3 实验结果

经过上一节的实验讨论，最后确定了 SIDL-SC 算法中各主要参数值如表2所示。图4展示了部分学习到的移不变基函数，在实验一中分别在数据集 A 中每类状态随机选择 200 个样本用于测试，实验结果如表 3 所示。

从表 3 可以看出，本文所提出的方法在数据集 A 中的表现良好，大部分样本都可以达到 100%的正确识别率，只有部分样本存在 1~2 个识别错误，随着数据集样本量的增大，本文算法可以取得更高的正确识别率。

表 2 SIDL-SC 各参数值

参数	数值
训练样本长度	512
基函数长度	80
稀疏化尺度	0.025
基函数个数	4
样本重叠率	50%
迭代次数	60

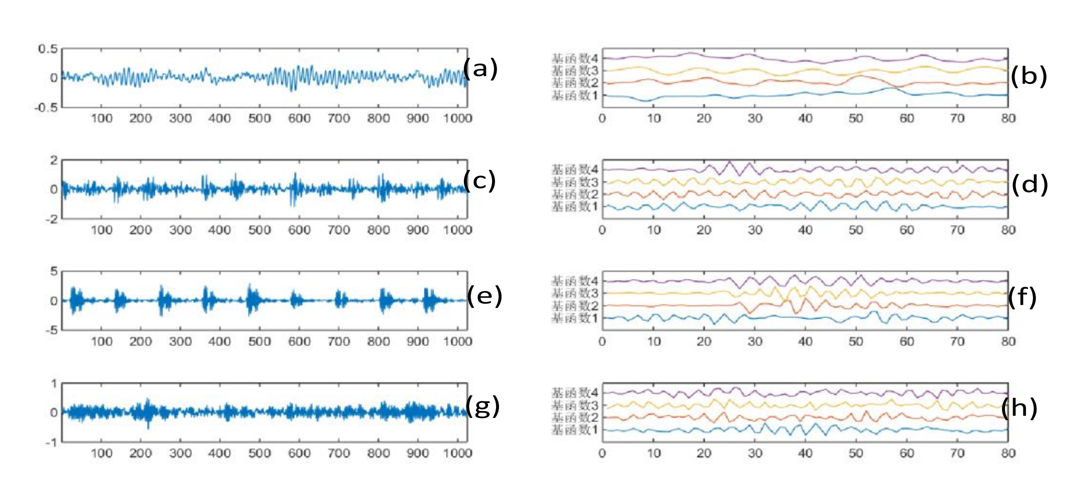


图 4　各状态典型振动信号及其对应的移不变基函数

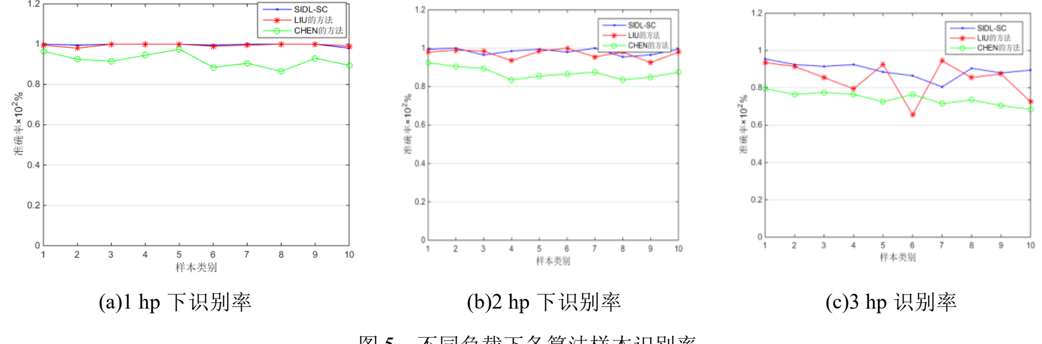
（a(b).正常状态；c(d).内圈故障；(f).外圈故障；g(h).滚动体故障）

表 3　数据集 A 实验结果

样本类别	识别正确率(%)	样本类别	识别正确率(%)
1	100	6	98.5
2	99.5	7	100
3	100	8	100
4	99	9	99.5
5	100	10	100

2.4　与传统方法对比实验

为了体现 SIDL-SC 算法在轴承故障识别中的优越性，本文与其他部分学者算法进行了对比实验。LIU[18]同样在字典学习和稀疏编码框架内实现对轴承故障的识别，其未探讨参数选择问题，且将所有基函数构成冗余字典，使待识别样本由该冗余字典重构，借助 LDA 分类器实现故障分类。CHEN[19]采集振动信号的标准差、偏度、峭度等时、频域特征送入卷积神经网络分类器，利用 LeNet5 模型实现对故障的识别。将以上两种方法和本文算法都在 0hp 下的样本进行训练，分别在数据集 B 中各负载下的样本作为测试，取 100 次实验的平均结果如图 5 所示。从实验结果可以看出，采用传统时频域特征用于分类的 CHEN 的算法具有相对较低的识别率，而 LIU 的算法虽然在个别样本上的识别率较高，但总体不如本文算法的稳定性高。



(a)1 hp 下识别率　　(b)2 hp 下识别率　　(c)3 hp 识别率

图 5　不同负载下各算法样本识别率

2.5　实测航空发动机实验

为验证本文算法在实际装备中的可行性，将该算法应用于实测某型航空发动机故障识别。实验数据采集于该型直升机涡轴发动机多台次、多次台架试车时所测得的四种状态下发动机振动信号，即正常状态、转子不平衡故障、转子松动故障和碰摩故障，采用等弧度采样方式，每周期采集 128 个数据点，典型振动信号如图 6 所示。

利用 SIDL-SC 算法对不同状态下的振动信号进行移不变字典学习，参数设置如下：训练样本长度为 256，基函数长度为 50，稀疏化尺度为 0.01，基函数个数为 2，其余参数设置与表 3 相同。学习到的移不变基函数如图 7 所示。

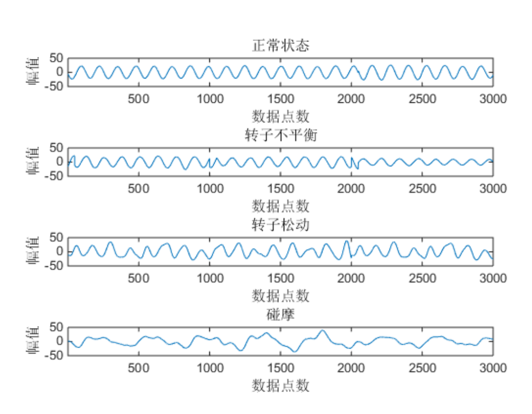
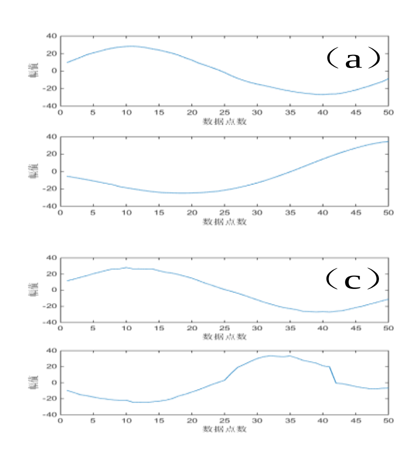


图 6　各状态典型振动信号



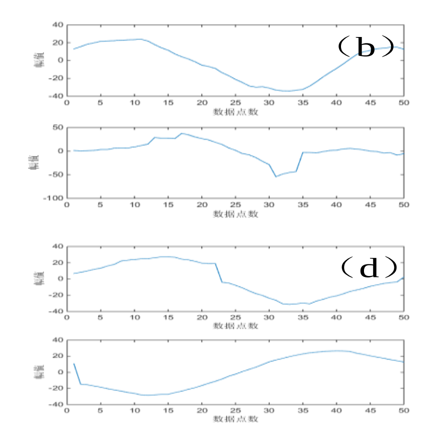


图 7　不同状态下学习到的移不变基函数（a.正常状态；b.转子不平衡；c.转子松动；d.碰摩）

从图中可以看出正常状态下的移不变基函数波形比较平缓，故障状态下学习到的移不变基函数均不同程度的含有其各自独特的特征。对不同状态进行识别，每组实验重复进行 100 次，取识别准确率作为评价指标，实验结果如表 4 所示，并与文献[18]和文献[19]中的方法进行对比。实验结果表明，由于发动机转子不平衡故障和转子松动故障与正常状态的波形上存在一些相似之处，导致算法的识别率低于 CWRU 数据集，但相较于上述两种算法，本文算法仍然能保持较高的故障识别率，进一步验证了本文算法的可行性。

表 4　航空发动机实验识别结果

识别率（%）	正常状态	转子不平衡	转子松动	碰摩
文献[18]方法	80	78	74	86
文献[19]方法	76	73	74	80
本文算法	85	82	81	93

3　结束语

SIDL-SC 是一种结合移不变字典学习和稀疏编码的振动信号故障识别算法，尤其是针对多工况条件下的故障识别有着较高的正确识别率。该方法以原始振动数据作为训练样本，通过特征自学习捕获反映不同状态信号本质特征的移不变基函数，以最小重构残差为准则判别故障的类别。同时对移不变字典学习中的惩罚因子进行了优化，以稀疏尺度作为前者的因变量，有效解决了不同训练样本长度下惩罚因子的设定问题。以 CWRU 振动数据库作为研究对象，本文详细探讨了不同参数选择对实验结果的影响，遴选出最优参数作为最终的实验参考变量。同时，将本文算法应用于实测航空发动机振动信号，并与其他部分学者的算法进行实验对比，结果表明，本文算法具有更强的泛化能力，可以更好的应用于实际振动数据的处理。

参考文献：

[1] 朱会杰，王新晴，芮挺，等. 基于频域信号的稀疏编码在机械故障诊断中的应用 [J]. 振动与冲击, 2015, 34 (12): 59-64.

[2] Janssens O, Slavkovikj V, Vervisch B, et al. Convolutional neural network based fault detection for rotating machinery [J]. Journal of Sound & Vibration, 2016, 377: 331-345.

[3] Harmouche J, Delpha C, Diallo D. Improved fault diagnosis of ball bearings based on the global spectrum of vibration signals [J]. IEEE Trans on Energy Conversion, 2015, 30 (1): 376-383.

[4] 王录雁，王强，张梅军，等. 基于 EMD 的滚动轴承故障灰色诊断方法 [J]. 振动与冲击, 2014, 33 (3): 197-202.

[5] 蔡剑华，胡惟文，王先春. 基于高阶统计量的滚动轴承故障诊断方法 [J]. 振动､测试与诊断, 2013, 33 (2): 298-301.

[6] Wright J, Yang A, Ganesh A, et al. Robust face recognition via sparse representation [J]. IEEE Trans on Pattern Analysis & Machine Intelligence, 2009, 31 (2): 210-227.

[7] 张新鹏，胡茑庆，程哲，等. 基于压缩感知的振动数据修复方法 [J]. 物理学报, 2014, 63 (20): 115-124.

[8] 郭亮，高宏力，黄海凤，等. 基于压缩感知理论的时变信号压缩技术 [J]. 西南交通大学学报, 2015, 50 (3): 511-516.

[9] 彭向东，张华，刘继忠. 基于过完备字典的体域网压缩感知心电重构 [J]. 自动化学报, 2014, 40 (7): 1421-1432.

[10] Aharon M, Elad M, Bruckstdin A M et al. K-SVD: an algorithm for designing overcomplete dictionaries for sparse representation [J]. IEEE Trans. on Signal Processing, 2006, 54 (11): 4311-4322.

[11] Smith E C, Lewicki M S. Efficient auditory coding [J]. Nature, 2006, 439 (7079): 978-982.

[12] Mørup M, Schmidt M N, Hansen L K. Shift invariant sparse coding of image and music data [R]. Technical Report, 2007.

[13] Tang Haifeng, Chen Jin, Dong Guangming. Sparse representation based latent components analysis for machinery weak fault detection [J]. Mechanical Systems & Signal Processing, 2014, 46 (2): 373-388.

[14] Grosse R, Raina R, Kwong H, et al. Shift-invariant sparse coding for audio classification [C]// Proc of Conference on Uncertainty in AI. Vancouver: AUAI Press, 2007: 149-158.

[15] Rubinstein R, Zibulevsky M, Elad M. Efficient implementation of the K-

SVD algorithm using batch orthogonal matching pursuit [J]. CS Technion, 2011, 40.

[16] Honglak L, Battle A, Raina R, et al. Efficient sparse coding algorithms [C]// Advances in Neural Information Processing Systems. 2006: 801-808.

[17] http://csegroups. case. edu/bearingdatacenter/pages/download-data-file [EB/OL].

[18] Liu Haining, Liu Chengliang, Huang Yixiang. Adaptive feature extraction using sparse coding for machinery fault diagnosis [J]. Mechanical Systems and Signal Processing, 2011, 25 (2): 558-574.

[19] Chen Zhiqiang, Li Chuan, Sanchez R V. Gearbox fault identification and classification with convolutional neural networks [J]. Shock & Vibration, 2015 (2): 1-10.